THE ABSTRACT OF THE DISCLOSURE

Please delete the original Abstract, and insert the following new Abstract:

The generation and management of a language model data structure include assigning each segment of a received corpus to a node in a data structure that denotes dependencies between the respective nodes. A transitional probability between each of the nodes in the data structure is calculated. A frequency of occurrence is calculated for each item of the respective segments, and those nodes of the data structure associated with items that do not meet a minimum frequency of occurrence threshold are removed. The data structure may be managed across a system memory of a computer system and an extended memory of the computer system.